1999

Bayes Estimation of a Distribution Function Using Ranked Set Samples

Paul H. Kvam
University of Richmond, pkvam@richmond.edu

Ram C. Tiwari

Bayes Estimation of a Distribution Function Using Ranked Set Samples

Paul H. Kvam, Georgia Institute of Technology

Ram C. Tiwari, University of North Carolina, Charlotte

Abstract

A ranked set sample (RSS), if not balanced, is simply a sample of independent order statistics generated from the same underlying distribution F. Kvam and Samaniego (1994) derived maximum likelihood estimates of F for a general RSS. In many applications, including some in the environmental sciences, prior information about F is available to supplement the data-based inference. In such cases, Bayes estimators should be considered for improved estimation. Bayes estimation (using the squared error loss function) of the unknown distribution function F is investigated with such samples. Additionally, the Bayes generalized maximum likelihood estimator (GMLE) is derived. An iterative scheme based on the EM Algorithm is used to produce the GMLE of F. For the case of squared error loss, simple solutions are uncommon, and a procedure to find the solution to the Bayes estimate using the Gibbs sampler is illustrated. The methods are illustrated with data from the Natural Environmental Research Council of Great Britain (1975), representing water discharge of floods on the Nidd River in Yorkshire, England.

Keywords

Dirichlet distribution, EM Algorithm, Gibbs sampler, Nonparametric estimation

1. Introduction

For populations in which the characteristic of interest can be easily identified and compared to other observations but not as easily quantified, ranked set sampling has proved to be a valuable tool of data acquisition. The procedure is a two-step sampling scheme in which a subgroup of independently sampled items are collected and ranked, but only one item from the subgroup is chosen for complete measurement. That item's rank within the subgroup is noted, so the final sample consists of independent order statistics. If subgroup sizes are identical (say k) and each of the k different order statistics are sampled in equal proportion, the ranked set sample is said to be balanced. As a sampling procedure, ranked set sampling achieved remarkable popularity in the applied statistical literature, as evidenced in the comprehensive review by Kaur, et al. (1995).

The ranked set sampling concept was introduced by McIntyre (1952) for application in an agricultural experiment in order to exploit the fact that measuring pasture yields was costly to the experimenter, while it was much easier for researchers in the field to examine the yields and rank them visually according to size (so long as k was reasonably small). Early research has concentrated on estimation of the unknown population mean. Takahasi and Wakimoto (1968) proved the assertion by McIntyre that the estimate of the population mean is improved using the sample average of a balanced ranked set sample (RSS) compared to the sample average of a simple random sample (SRS) of the same size. It was also shown that the RSS mean is the only unbiased linear estimator of the population mean, and the variances of the two estimators, denoted σ^2_{SRS} and σ^2_{RSS}, satisfy the inequality $\frac{2}{k+1}\sigma^2_{SRS} \le \sigma^2_{RSS} \le \sigma^2_{SRS}$. Efficiency comparisons for particular parametric families of interest are investigated by Dell and Clutter (1972).

The scope of inference under RSS was broadened by Stokes and Sager (1988) to estimation of the underlying distribution function, F. They showed that the empirical distribution function (EDF) based on a RSS is unbiased for F and is uniformly more efficient in estimating F than the EDF for a SRS of the same size. More recently, Kvam and Samaniego (1994) derived a nonparametric maximum likelihood estimator (MLE) for F that no longer required a balanced RSS; that is, any collection of independent order statistics can be used to estimate F.

In many suitable applications, we might expect that supplemental prior information exists regarding the unknown distribution. In such applications, Bayes methods have intuitive appeal for both parametric and nonparametric modeling. For instance, there might be past data available that can be quantified into a prior distribution. Chapter 3 of Berger (1985) elucidates methods of quantifying such prior information in an objective way. We investigate the Bayesian counterpart of the estimation problem for ranked set samples, emphasizing the problem of estimating the unknown distribution function.

In Section 2, we discuss the derivation of the posterior distribution for F, based on a singular ordered Dirichlet distribution as the prior. The Bayes estimator is derived for squared error loss, and the Bayes generalized maximum likelihood estimator (GMLE), described by DeGroot (1970), is also obtained. In general, no closed form solutions for estimates of F exist in either case, thus iterative techniques are discussed to obtain them. Section 2.1 displays an iterative procedure, based on the EM Algorithm, to find GMLEs of F. In section 2.2, the standard Bayes estimator with squared error loss is considered. In this case, a simple solution is not apparent. We present an iterative procedure based on the Gibbs sampler. Both techniques are illustrated in Section 3, using water discharge data from floods of the Nidd River in Yorkshire, England during the years 1963 to 1968.

2. Nonparametric Bayes Estimation

In this section, we formulate a Bayes estimator for F by incorporating prior information along with the likelihood of the observed RSS data. Let X_i represent the ith item from the RSS, i = 1,...,n. Corresponding to the subsample from which the item was measured, we denote the subsample size and the rank of X_i within the subsample as k_i and r_i, respectively, thus $1 \le r_i \le k_i$. Some of the work that follows extends the results of Kvam and Samaniego (1994), referred to as K&S from this point on. In order to be consistent with the notation in K&S, we further denote the ordered sample as $X_{(1)} < \ldots < X_{(n)}$, and let $k_{(i)}$ and $r_{(i)}$ correspond to the ranked value $X_{(i)}$. In this case, $k_{(i)}$ and $r_{(i)}$ are concomitants of $X_{(i)}$, and are not necessarily ordered themselves. If F is absolutely continuous, the density of $X_{(i)}$ is expressed in terms of the density function:

$$f_i(x) = r_i \binom{k_i}{r_i} F(x)^{r_i - 1}(1 - F(x))^{k_i - r_i} f(x). \qquad (1)$$

If F is not absolutely continuous, we cannot assume that the density in (1) exists. Instead, we take an approach along the lines of Kiefer and Wolfowitz (1956) in modeling the likelihood of the data. From (1), the likelihood function for the observed RSS is

$$h(x_1,\ldots,x_n \mid F) \propto \prod_{i=1}^{n} F(x_{(i)})^{r_{(i)}-1}[1-F(x_{(i)})]^{k_{(i)}-r_{(i)}} dF(x_{(i)}). \tag{2}$$

where dF(x) represents the difference F(x) - F(x-). It can be seen that h = 0 if any point x_i is not assigned positive mass. Furthermore, distributions that assign mass to points or intervals outside of the set of observed values cannot maximize the likelihood function (see, for instance, Miller (1981) for further discussion on the nonparametric likelihood problem). This fact has significance that will be exploited in Section 2.1. The estimation of F is now reduced to that of the (aggregated) multinomial parameters $\phi_i = F(x_{(i)})$, i = 1,...,n. Note that the distribution F is characterized by the vector $\phi = (\phi_1,\ldots,\phi_n)$.

To model uncertainty of the experimenter's belief about the distribution F, Ferguson (1973) introduced a Dirichlet process prior which serves as a standard prior distribution for nonparametric estimation problems. For the RSS setting, we adopt a joint ordered Dirichlet distribution for the parameters $(\phi_1,\ldots,\phi_n)$, denoted $D(\alpha_1,\ldots,\alpha_{n+1})$. This distribution serves as a conjugate prior for F in a simple random sample, thus it is a natural selection as a prior distribution for this problem even though it is not conjugate with a general RSS. The joint prior distribution of $\phi = (\phi_1,\ldots,\phi_n)$, with $\phi_0 = 0$, is expressed as

$$g(\phi \mid \alpha_1,\ldots,\alpha_n) \propto \left(\prod_{i=1}^{n} (\phi_i - \phi_{i-1})^{\alpha_i - 1}\right)(1-\phi_n)^{\alpha_{n+1}-1}. \tag{3}$$

over the space $\Phi = \{ \phi : 0 \le \phi_1 \le \ldots \le \phi_n \le 1 \}$.

The ordered Dirichlet serves as a flexible tool for modeling prior information. Any distribution can be used as a prior estimate of F, and the uncertainty involved with this guess is reflected in the sum of its n+1 parameters, A = $\sum_{j=1}^{n+1} \alpha_j$, which can be set at the user's discretion. To see how the prior can be constructed, suppose the subjective information suggests that F_0 is the best guess for the unknown distribution F. Let $p_i = F_0(x_{(i)}) - F_0(x_{(i-1)})$, i=1,...,n+1, allowing $x_{(0)} = -\infty$ and $x_{(n+1)} = \infty$. We assign weight to the guess F_0 by equating its certainty with an amount of data, say m ≥ 0. The ordered Dirichlet is characterized by (F_0, m, n), or equivalently, $(\alpha_1, ..., \alpha_{n+1})$ where $\alpha_i = (m+n+1)p_i$, i=1,...,n+1, and A = m + n + 1.

The posterior distribution for ϕ is not an ordered Dirichlet. If we denote α = $\{\alpha_1, ..., \alpha_{n+1}\}$, the posterior density function can be expressed as

$$g(\phi \mid x_1, ..., x_n, \alpha) . \propto . \prod_{i=1}^{n} \phi_i^{r_{(i)}-1}(1-\phi_i)^{\beta_i}(\phi_i - \phi_{i-1})^{\alpha_i}, \quad \phi \in \Phi, \qquad (4)$$

where $\beta_i = \begin{cases} k_{(i)} - r_{(i)} & i = 1, ..., n-1 \\ k_{(n)} - r_{(n)} + \alpha_{(n+1)} - 1 & i = n \end{cases}$.

Let us now consider two alternative estimators of F using the Bayesian paradigm: the GMLE and the Bayes estimator with respect to the squared-error loss function.

2.1 The Generalized Maximum Likelihood Estimator

Here, the GMLE is the value of F that satisfies

$$g(\phi^* \mid x_1,\ldots,x_n,\alpha) = \max_{\phi \in \Phi} g(\phi \mid x_1,\ldots,x_n,\alpha).$$

The estimator is not necessarily a Bayes estimator with respect to a standard loss function. However, the derivation of this GMLE parallels maximum likelihood theory, thus giving the estimate an easier frequentist interpretation. In this case, we simply maximize (4) with respect to ϕ, which is equivalent to solving the following likelihood equations:

$$\frac{\partial}{\partial\phi_1}\ln g(\phi \mid x_1,\ldots,x_n,\alpha) = \frac{r_{(1)}+\alpha_1-1}{\phi_1} - \frac{k_{(1)}-r_{(1)}}{1-\phi_1} - \frac{\alpha_2}{\phi_2-\phi_1} = 0 \qquad (5)$$

$$\frac{\partial}{\partial\phi_i}\ln g(\phi \mid x_1,\ldots,x_n,\alpha) = \frac{r_{(i)}-1}{\phi_i} - \frac{k_{(i)}-r_{(i)}}{1-\phi_i} - \frac{\alpha_{i+1}}{\phi_{i+1}-\phi_i} + \frac{\alpha_i}{\phi_i-\phi_{i-1}} = 0$$

for i = 2,...,n-1, and

$$\frac{\partial}{\partial\phi_n}\ln g(\phi \mid x_1,\ldots,x_n,\alpha) = \frac{r_{(n)}-1}{\phi_n} - \frac{k_{(n)}-r_{(n)}+\alpha_{n+1}-1}{1-\phi_n} + \frac{\alpha_n}{\phi_n-\phi_{n-1}} = 0.$$

The posterior distribution, as expressed in (4), shares several attributes with the likelihood function found in K&S. Consequently, some properties of the nonparametric MLE can be derived for the GMLE. We use the proofs found in the former paper when proposing extensions for this Bayes estimator, but for simplicity, redundant portions of analogous proofs will be left out or abbreviated in this article. For instance, the Hessian matrix corresponding to the log-posterior distribution can be easily shown to be non-positive definite using a simple extension of Theorem 2.1 of K&S, thus the following result holds for any vector $\alpha \geq (0,...,0)'$:

Theorem 1. Using an ordered Dirichlet prior D($\alpha_1,...,\alpha_{n+1}$), the solution ($\phi$*) to the equations in (5) is the unique Bayes GMLE for F.

Note: If $r_{(1)}$=0 and $\alpha_1 \leq 1$, the likelihood is maximized at ϕ_1*=0. This is an intuitive solution, because $r_{(1)}$=0 means that the earliest observation represents a subsample minima, thus no data value less than this can be confirmed. Analogously, if $r_{(n)}=k_{(n)}$, meaning that the last observation represents a subsample maxima, and if $\alpha_{n+1} \leq 1$, then the likelihood is maximized at ϕ_n*=0.

In general, a closed form solution to ϕ* is not available. We now discuss an iterative procedure used to create a sequence of estimators {$\phi^{(0)}$, $\phi^{(1)}$, $\phi^{(2)}$, ...} that converges to ϕ*. The Dirichlet prior for ϕ leads to a practical choice for $\phi^{(0)}$: $\phi_1^{(0)} = \frac{\alpha_1}{A}$, $\phi_2^{(0)} = \frac{\alpha_1 + \alpha_2}{A}$, ..., $\phi_n^{(0)} = \frac{\alpha_1 + ... + \alpha_n}{A}$. In this case, $\phi_i^{(0)} - \phi_{i-1}^{(0)} = \frac{\alpha_i}{A}$, and $\phi^{(0)} \in \Phi$.

Theorem 2. With an initial estimate of $\phi = \phi^{(0)}$, we update the estimate in the jth step (for i = 1,...,n) from $\phi^{(j)}$ to

$$\phi_i^{(j+1)} = \frac{\sum_{s=1}^{i}\left(r_{(s)}+\alpha_s-1\right)+\sum_{s=1}^{i-1}(k_{(s)}-r_{(s)})\frac{\phi_i^{(j)}-\phi_s^{(j)}}{1-\phi_s^{(j)}}+\sum_{t=i+1}^{n}\left(r_{(t)}-1\right)\frac{\phi_i^{(j)}}{\phi_s^{(j)}}}{\sum_{l=1}^{n}k_{(l)}+\sum_{l=1}^{n+1}\left(\alpha_l-1\right)}. \tag{6}$$

Sketch of Proof: The updating equations in (6) are a simple Bayes extension of the EM Algorithm as described by Dempster, Laird and Rubin (1977). This iterative procedure is a practical choice because the posterior distribution is maximized to find the GMLE for ϕ. For the case in which α equals the vector of ones, the solution is exactly the MLE, and the iterative equation (6) has the following intuition: if we had a complete sample of $\sum_{l=1}^{n} k_{(l)}$ observations (not just the order statistics), then $\phi_i^{(j+1)}\sum_{l=1}^{n} k_{(l)}$ approximates the number of iid observations less than or equal to $x_{(i)}$. Using our previous estimate of F ($\phi^{(j)}$), then for any actual observation $x_{(s)} < x_{(i)}$, we know $r_{(s)}$ of $k_{(s)}$ from that subsample must be less than $x_{(i)}$. Of the $(k_{(s)}-r_{(s)})$ other unobserved items in the subsample, we estimate that each one has a probability equal to $\frac{\phi_i^{(j)}-\phi_s^{(j)}}{1-\phi_s^{(j)}}$ of being less than $x_{(i)}$. Furthermore, for actual observations $x_{(t)} > x_{(i)}$, it is not clear that any of the $k_{(t)}$ are less than $x_{(i)}$, but of the $\left(r_{(t)}-1\right)$ unobserved items that have such a chance, we estimate each one has a probability equal to $\frac{\phi_i^{(j)}}{\phi_s^{(j)}}$ of being less than $x_{(i)}$. From this, we see that (6) updates the estimate of ϕ by estimating the unobserved observations and formulating the EDF based on the hypothetical (complete) data set of size $\sum_{l=1}^{n} k_{(l)}$.

The Bayes extension can be interpreted using the same intuition, but first adding weights $\alpha_s - 1$ to sth smallest observation. In this case, the complete sample size increases to $\sum_{l=1}^{n} k_{(l)} + \sum_{l=1}^{n+1} (\alpha_l - 1)$. This shows how the updating equation in (6) is a simple extension of the iterations from (15) and (17) of K&S which implies the iterations are a form of the EM Algorithm. □

Conditions for convergence for the EM Algorithm were over extended by Dempster, Laird and Rubin (1977). Both Boyles (1983) and Wu (1983) prove that convergence cannot be guaranteed in the generality claimed in that paper. However, we note that g(ϕ | $x_1,...,x_n, \alpha$), as a function of the data, is bounded and proportional to some density belonging to the exponential family, From this fact, the continuity conditions of Wu (1983) are satisfied, which implies that the solution to the iterations in (6) converges to a stationary point. If there exists a mapping from the equations pertaining to the solution to this stationary point to the solution for (5), the stationary point must be the Bayes Estimator from Theorem 1. This mapping is depicted in the Appendix, thus the following holds:

Theorem 3. The sequence of estimators in (6) converges to the Bayes GMLE for F.

In some examples, the algorithm is slow to converge. For the data illustrated in Section 3, as example, ten iterations are required before the answer is stable past two decimal points, and over 40 are required before the solution is constant in four decimal places.

2.2 Bayes Estimation Using Squared Error Loss

In this section, we investigate the Bayes estimate of F with respect to squared error loss. Extensions to other common loss functions, such as absolute-error loss, will be apparent to the reader. Typical of problems involving RSS-based distributions, the expected value of the posterior distribution has no simple analytical derivation. If the squared error loss function (that leads us to find the Bayes estimator of F as the expected value) is applied, numerical methods are required to find the solution. The Gibbs sampler is used below to formulate Bayes estimates in this case. Compared to the GMLE, estimates based on squared error loss require more numerical computation, and the formulation of the estimator is more complicated.

The Gibbs sampler of Geman and Geman (1984) is a procedure for generating random samples from a multidimensional distribution without having to calculate the exact distribution of origin. Full conditional distributions, assumed to be relatively easier to handle than the complete (marginal) distribution, are used to generate a random sample from the posterior density. From this sample, the mean can be used as the Bayes estimate of F.

To implement the Gibbs sampling technique, the conditional distributions of ϕ_i given $(\phi_1, ..., \phi_{i-1}, \phi_{i+1}, ..., \phi_n)$ are required. Based on the posterior density in (4), we have

$$g_1(\phi_1 \mid \phi_2, ..., \phi_n, data) .\propto. \ \phi_1^{r_{(1)} + \alpha_1 - 1}(1-\phi_1)^{\beta_1}(\phi_2 - \phi_1)^{\alpha_2}, \quad 0 < \phi_1 < \phi_2 \ ; \qquad (7)$$

$$g_2(\phi_2 \mid \phi_1, \phi_3, ..., \phi_n, data) .\propto. \ \phi_2^{r_{(2)} - 1}(1-\phi_2)^{\beta_2}(\phi_2 - \phi_1)^{\alpha_2}(\phi_3 - \phi_2)^{\alpha_3}, \quad \phi_1 < \phi_2 < \phi_3;$$

$$\vdots$$

$$g_n(\phi_n \mid \phi_1,...,\phi_{n-1}, data) . \propto . \phi_n^{r_{(n)}-1}(1-\phi_n)^{\beta_n}(\phi_n - \phi_{n-1})^{\alpha_n}, \ \phi_{n-1} < \phi_n \le 1.$$

Although the multivariate distribution for ϕ is untenable in this estimation problem, the (univariate) conditional distributions provide less difficulty in generating random samples. We start with the same initial guess $\phi^{(0)}$ from Section 2.1, then sample $\phi^{(1)}$ by generating $\phi_i^{(1)}$ from the conditional distribution of $\phi_i \mid \phi_1^{(1)},...,\phi_{i-1}^{(1)},\phi_{i+1}^{(0)},...,\phi_n^{(0)}$ as i increases from 1 to n. After k iteration of this procedure, we have $\phi^{(k)} = (\phi_1^{(k)},...,\phi_n^{(k)})$. If k is large enough, $\phi^{(k)}$ is regarded as a sample from the posterior distribution mentioned in (4). In choosing the value of k for which the Gibbs sequence can be concluded, a simple check based on differences in consecutive values in the sequence should suffice. However, Casella and George (1992) discuss how this approach cannot be guaranteed. In many applied problems, k = 40 should ensure convergence.

This represents a single (multivariate) observation for the Gibbs sample. To obtain a reasonable estimate of the posterior distribution mean, we repeat this iterative procedure N times, usually with N much larger than k, such as N = 2000. Let $\phi(j)$ represent the jth independent replication of ϕ from the Gibbs procedure, with the superscript (k) suppressed. Then $\phi(j) = (\phi_1(j),...,\phi_n(j))$, and the Bayes estimate of ϕ is the Gibbs sample mean:

$$\hat{\phi}_i = \frac{1}{N}\sum_{j=1}^{N}\phi_i(j), \quad i = 1,...,n. \tag{8}$$

One can obtain the posterior median using the Gibbs sampler, thus the Bayes estimator with respect to absolute-error loss can be obtained in the same manner, replacing (8) with $\tilde{\phi}_i$ = Median($\phi_i(j)$, j = 1,...,N) ,   i = 1,...,n.

An approximate (1-p)100% credible set for the vector ϕ can be constructed by selecting the 100(p/2) and 100(1-p/2) percentiles of the Gibbs sample for each value of ϕ_i , i= 1,...,n. For example, an upper 95% credible interval for F is calculated as

$\hat{\phi}_i^{(0.95)}$ = 95th percentile of the set { ϕ_i (1),..., ϕ_i (N)}, i = 1,...,n.

Finding Bayes estimates this way presents more computational problems than found in Section 2.1. The crux of the problem lies in generating values from the conditional distributions in (7). Some modern statistical software can be used to generate values (we used Mathematica by Wolfram Research, Inc.), and Fortran users can generate random variables from general distributions using the IMSL library.

To generate values from the conditional distributions in (7), the rejection/acceptance algorithm described by Tanner (1992) can be implemented by first finding a set of simple functions $(\gamma_1,...,\gamma_n)$ that cover the densities in (7). In this case, if $\gamma_i(x) \geq g_i(x)$, γ_i is not necessarily a density function, thus we define a probability density function (pdf) $q_i(x) = \gamma_i(x)/c$, where c = $\int_0^\infty \gamma_i(u)du$. The algorithm has two basic steps: first generate a random value from the simple pdf $q_i(x)$ along with a random variable u from U(0,1). If u $\leq g_i(x)/\gamma_i(x)$, we keep the value from $q_i(x)$, otherwise we repeat the procedure. The obtained set represents random observations from the conditional densities in (7).

3. Example

Table 1 lists data from the Natural Environmental Research Council of Great Britain (1975). The data values represent water discharges (in cubic meters per second) on the Nidd River in Yorkshire, England, for the first observed flood of the year during 1963 to 1968. In this case, only the first three floods in those years are considered, and the sample of six flood records comprises a balanced ranked set sample. However, this is a matter of random coincidence more than planned sampling. This sample illustrates how ranked set sampling can be economical. In this case, only the first flood of three observed floods needs to be measured. The remaining two floods are tabulated and it is noted whether these floods surpass the first one in terms of water discharge.

Table 1 displays the Bayes GMLE and the Bayes estimates using both absolute-error loss and squared-error loss functions with the Nidd River data. For this example, we inserted a non-informative prior of $\alpha_i = 1$, $i \geq 1$, and the table suggests that all three estimators are similar for this example. A plot of the squared error lost estimate of F appears in Figure 1. The dotted lines represent the 90% credible interval for the estimate, computed using the numerical algorithm outlined in Section 2.2. A stream of k = 40 values were generated for each Gibbs sequence, and the iterative process was repeated N = 2000 times.

The Gibbs sampler can be used to obtain the GMLE, but at greater computational cost. Furthermore, modal values for samples of continuous data are less reliable than other estimators of population centrality. Figure 2 summarizes the Gibbs sample obtained for approximating values for Figure 1, and the modal values of $(\phi_1,...,\phi_n)$ can be inferred.

In the last example, if stronger prior information is supplied for estimating the river's water discharge during a flood, the resulting estimator can change dramatically. To illustrate the potential difference, we construct the Bayes GMLE again, but this time based on a normal distribution with mean 80 cm/s, standard deviation 20 cm/s, and prior distribution weight A = 27. This corresponds to expert opinion that dictates the water discharge has a Normal distribution (μ=80, σ=20), and is weighted to be equivalent to m = A - n - 1 = 20 observations. By the methods described in Section 2, the prior is characterized by the vector $\alpha_i = A(\Phi(x_{(i)}) - \Phi(x_{(i-1)}))$, i = 1,...,7. Figure 2 shows how the prior distribution dominates the empirical evidence for such an example. The new GMLE represents a compromise between the prior distribution function and the GMLE based on the noninformative prior.

References

Berger, J. O. (1985) *Statistical Decision Theory and Bayesian Analysis*, Springer-Verlag, New York.

Casella, G. and George, E. I. (1992), “Explaining the Gibbs sampler,” *The American Statistician*, Vol. 46, No. 3, 167 - 174.

DeGroot, M. (1970), *Optimal Statistical Decisions*, McGraw-Hill, New York.

Dell, T. R. and Clutter, J. L. (1972), “Ranked set sampling theory with order statistics background.” *Biometrics*, 28, 545 - 553.

Dempster, A. P., Laird, A., and Rubin, D. B. (1977), "Maximum likelihood from incomplete data via the EM Algorithm," *Journal of the Royal Statistical Society*, *Ser. B*, 39, 1-38.

Ferguson, T. S. (1973), “A Bayesian analysis of some nonparametric problems,” *Annals of Statistics*, 1, 209 - 230.

Geman, S. and Geman D. (1984), “Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6, 721 - 741.

Kaur, A., Patil, G. P., Sinha, A. K. and Taillie, C. (1995), “Ranked set sampling: an annotated bibliography,” *Environmental and Ecological Statistics*, 2, 25 - 54

Kiefer, J. and Wolfowitz, J. (1956), “Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters.” *The Annals of Mathematical Statistics*, 27, 887 - 906.

Kvam, P. H. and Samaniego, F. S., “Nonparametric maximum likelihood estimation based on ranked set samples.” *Journal of the American Statistical Association*, 89, No. 426, 526 - 537.

McIntyre, G. A. (1952), “A method of unbiased selective sampling using ranked sets.” *Australian Journal of Agricultural Research*, 3, 385 - 390

Miller, R. (1981), *Survival Analysis*, John Wiley and Sons, Inc. New York.

Natural Environmental Research Council of Great Britain (1975), *Flood Studies Report*, Volume 4, 235 - 236. London: Whitefriars Press.

Takahasi, K. and Wakimoto, K. (1968), “On unbiased estimates of the population mean based on samples stratified by means of ordering.” *Annals of the Institute of Statistical Mathematics*, 20, 1 - 31.

Tanner, M. A. (1992), *Tools for Statistical Inference*, Lecture Notes in Statistics, Vol. 67, Springer-Verlag, New York.

Wu, C. F. J. (1983) "On convergence properties of the EM algorithm," *Annals of Statistics*, 11, 95-103.

Appendix

To prove the solution to the iterative scheme in (6) solves the likelihood equations, it is shown below that the self-consistency equations that characterize the unique solution of (6) are equivalent to the estimating equations of the log-likelihood in (5). Let $K^+ = \sum_{l=1}^{n} k_{(l)} + \sum_{l=1}^{n+1}(\alpha_l - 1)$. The solution corresponding to the self consistency equations can be expressed as

$$\phi_i^* = \frac{\sum_{s=1}^{i}\left(r_{(s)} + \alpha_s - 1\right) + \sum_{s=1}^{i-1}(k_{(s)} - r_{(s)})\frac{\phi_i^* - \phi_s^*}{1-\phi_s^*} + \sum_{t=i+1}^{n}\left(r_{(t)} - 1\right)\frac{\phi_i^*}{\phi_s^*}}{\sum_{l=1}^{n} k_{(l)} + \sum_{l=1}^{n+1}\left(\alpha_s - 1\right)}, \quad i = 1,...,n. \qquad \text{(A1)}$$

By taking differences $\phi_i^* - \phi_{i-1}^*$ in (A1) and dividing each difference by the same, we have n equations $M_i(\phi^*) = 0$, where

$$M_i(\phi^*) = K^+ - \left(\sum_{j=i}^{n}\frac{r_j - 1}{\phi_j^*} + \sum_{j=1}^{i-1}\frac{k_j - r_j}{1-\phi_j^*} - \frac{\alpha_i}{\phi_i^* - \phi_{i-1}^*}\right). \qquad \text{(A2)}$$

Taking successive differences $M_i(\phi^*) - M_{i+1}(\phi^*)=0$ over i = 1,...,n-1 produces the first n-1 estimating equations of the log-likelihood. The final estimating equation is realized from the identity $M_n(\phi^*)=0$.

$x_{(i)}$	$k_{(i)}$	$r_{(i)}$	ϕ_i (GMLE)	ϕ_i (SEL)	ϕ_i (AEL)
80.12	3	3	0.1573	0.2361	0.2304
87.76	3	2	0.3386	0.3650	0.3595
99.08	3	3	0.4947	0.4684	0.4659
111.54	3	1	0.5903	0.5641	0.5680
121.73	3	1	0.7699	0.6723	0.6793
123.71	3	2	0.8923	0.8454	0.8590

Table 1. Flood Discharge levels, the GMLE and Bayes Estimate for squared error loss (SEL) and absolute-error loss (AEL).

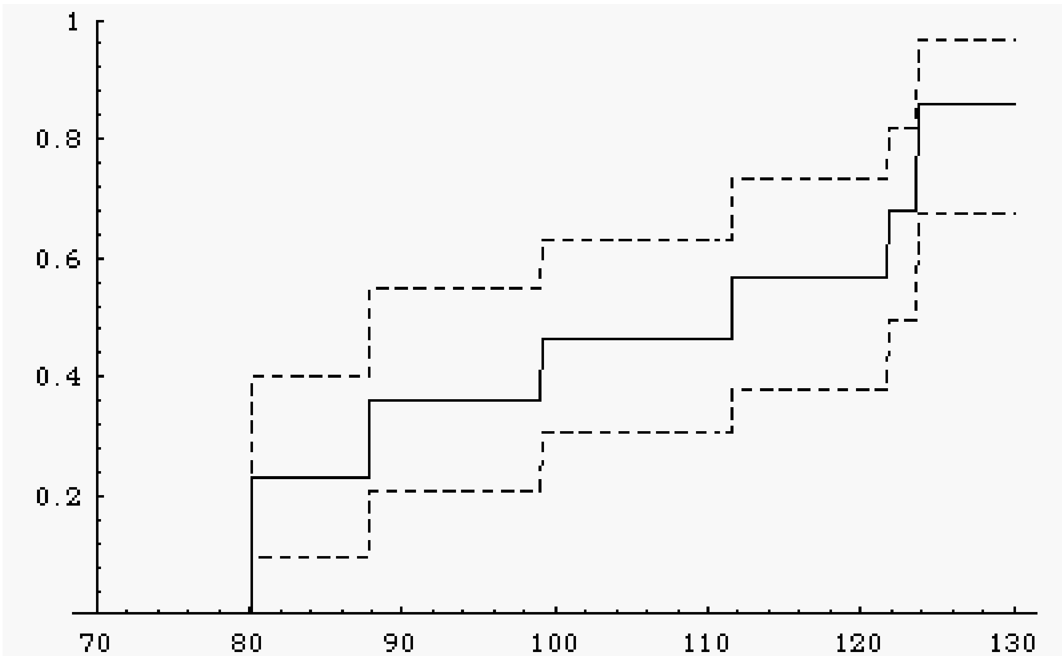

Figure 1. Squared-error loss estimate of F along with corresponding 90% credible interval using Nidd River flood data and non-informative prior

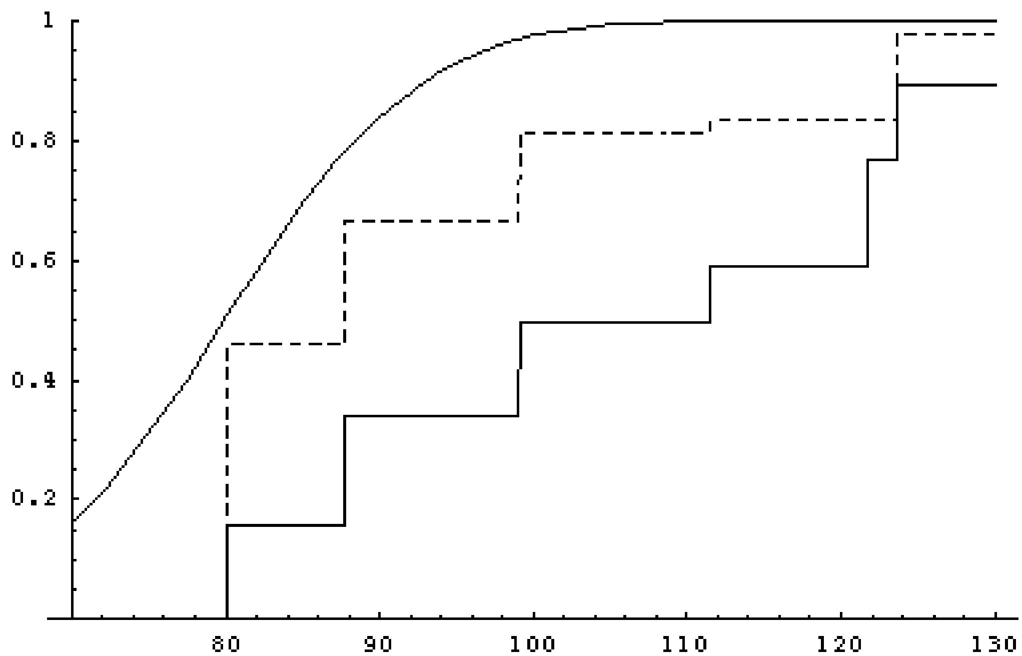

Figure 2 . GMLE based on noninformative prior (solid line), GMLE based on prior based on normal distribution (dashed line) and the distribution function for the prior.

Biographical Sketch

Paul H. Kvam is an assistant professor in the school of Industrial and Systems Engineering at Georgia Institute of Technology. He received his B.S. in Mathematics from Iowa State University in 1984, M.S. in Statistics from the University of Florida in 1986, and Ph.D. from the University of California, Davis in 1991. His research areas include statistical reliability with applications to engineering problems. From 1991 to 1995, Dr. Kvam was a staff scientist at Los Alamos National Lab in New Mexico, where he worked on sampling and estimation problems related to environmental restoration and toxic clean-up at various remote laboratory sites. He is a member of ASA and IMS.

Ram C. Tiwari is Professor, Department of Mathematics, University of North Carolina at Charlotte, Charlotte, NC 28269. He has published more than 100 papers in many international journals on topics such as Bayesian nonparametric inference, survival analysis, reliability theory, and Kalman filtering. He is a member of ASA and IMS.

Acknowledgment

We are grateful to the editor and the referees for their helpful comments and suggestions, which led to several improvements in the manuscript .